

Commercialized Language Training and Phonological Fossilization

(A Case Study of Non-Standard Romanization in Indonesian Migrant Workers' Mandarin Curricula)

Arjuna Satria Dewa Bagaskara^{1*}, Ressi Maulidina Delijar²

¹ Faculty of Cultural Studies, Universitas Brawijaya, Indonesia
e-mail: dewaagaskara@student.ub.ac.id

² Faculty of Cultural Studies, Universitas Brawijaya, Indonesia
e-mail: ressimd@ub.ac.id

* Corresponding Author: dewaagaskara@student.ub.ac.id

Abstract. Pre-departure Mandarin training for Indonesian migrant workers (PMI) in Taiwan is predominantly delivered through private job training institutions (LPKS) operating under commercial pressure and abbreviated timelines. This study examines how a non-standard, L1-based romanization system used in one such institution shapes the acoustic production of Mandarin affricates in five female PMI currently employed in Taiwan. An exploratory mixed-methods case study design was employed, triangulating acoustic phonetic analysis of /z/, /c/, /j/, /q/, /zh/, and /ch/ through Voice Onset Time (VOT), Frication Duration (FD), and Center of Gravity (CoG) measurements with thematic analysis of semi-structured interviews and systematic document analysis of the training module. Findings reveal pervasive over-aspiration of non-aspirated affricates, collapse of aspiration contrasts in the alveolar and palatal series, and systematic dentalization of retroflex affricates attributable to CoG values substantially exceeding native ranges. Triangulation across data sources demonstrates that these deviations are traceable to specific grapheme-phoneme mappings in the training module that activate Indonesian L1 phonological categories in place of Mandarin targets. Fossilization is reinforced by a workplace ecology that provides no corrective feedback. The study argues that these outcomes constitute a systemic curriculum failure and calls for regulatory intervention by Kemnaker and BP2MI to standardize phonological content in LPKS language teaching materials.

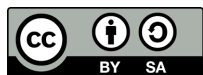
Keywords: Curriculum Policy; Indonesian Migrant Workers; Mandarin Affricates; Phonological Fossilization; Romanization.

1. Introduction

Every year, hundreds of thousands of Indonesian citizens depart for Taiwan as migrant workers. Most of them carrying a working proficiency in Mandarin that was built almost entirely inside a short, intensive training program at a private job training institution. These institutions, known in Indonesia as Lembaga Pelatihan Kerja Swasta (LPKS), operate under considerable commercial and time pressure, and the language instruction they deliver reflects that pressure in ways that are rarely examined with empirical rigor (BP2MI, 2025; Purba et al., 2020). What gets compressed or distorted under these conditions is frequently phonological instruction, the aspect of language learning that most directly determines whether a worker will be understood in the daily demands of a Mandarin-speaking workplace.

This study began with a small but consequential observation. In the Mandarin training module produced by LPKS Eka Management Cipta Sejahtera in Ponorogo, East Java, certain Mandarin words were transcribed not in standard Hanyu Pinyin but in locally adapted spellings drawn from Indonesian orthographic conventions. The word *měi* appeared as “Mei Jhe,” and *zǐ* was printed as “Ce Ci” (Afrina & Cleveresty, 2021; Bagaskara, 2026; LPKS, n.d.). For a learner whose phonological world is shaped by Indonesian, these spellings are not neutral decoding aids. They are active invitations to replace Mandarin phonological categories with familiar L1 ones, and the consequences of that invitation are traceable, with acoustic precision, in the speech of the workers who learned from them.

Received: April 23, 2026
Revised: May, 12, 2026
Accepted: June 24, 2026
Published: July 01, 2026
Curr. Ver.: July 01, 2026



Copyright: © 2025 by the authors.
Submitted for possible open
access publication under the
terms and conditions of the
Creative Commons Attribution
(CC BY SA) license
(<https://creativecommons.org/licenses/by-sa/4.0/>)

The Mandarin affricate system is among the most demanding acquisition targets for speakers of languages that lack both aspirated consonants and retroflex articulations. Mandarin draws phonemic contrasts between aspirated and unaspirated affricates across alveolar, palatal, and retroflex places of articulation, contrasts cued primarily by Voice Onset Time (VOT), Frication Duration (FD), and Center of Gravity (CoG) in the spectral domain (Lee-Kim, 2014; S. Li & Gu, 2015). Indonesian operates with a voiceless-voiced plosive system in which aspiration carries no phonemic weight and retroflex consonants are entirely absent from the inventory (Alwi et al., 2000; H. Li et al., 2019). Adult Indonesian learners attempting to acquire Mandarin affricates carry this phonological configuration with them, and the existing literature documents systematic errors in aspiration timing and place of articulation as predictable consequences (Fitriyah & Subandi, 2020; D. Liu & Reed, 2021).

What has received far less attention is the role of the romanization system itself as a variable that either supports or undermines phonological category formation. Prior studies on Indonesian learners of Mandarin have focused predominantly on university students with access to standard Pinyin and professional instruction (Fitriyah & Subandi, 2020; H. Li et al., 2019; D. D. Liu et al., 2021). Research on orthographic effects in L2 Chinese phonology has similarly centered on tone production or vowel phenomena in formal educational settings (Bassetti, 2007; J. Li et al., 2014; X.-H. Li et al., 2011). Studies of PMI in Taiwan have generally addressed sociolinguistic positioning or communicative strategies rather than the acoustic consequences of training materials (Fauzi, 2021; Thamrin et al., 2022; Widodo & Belgradoputra, 2013). The present study is positioned at the intersection of these bodies of work, where empirical examination has not yet been carried out with the care the question deserves.

Two objectives are pursued. The acoustic characteristics of PMI affricate production are described through VOT, FD, and CoG measurements compared against native reference values. The pedagogical, phonological, and psycholinguistic factors underlying those characteristics are analyzed and connected to questions of training policy and the linguistic rights of migrant workers. They are more accurately understood as the foreseeable output of an institutional system in which phonological accuracy was subordinated to commercial efficiency, and in which no regulatory mechanism was put in place to prevent that subordination from occurring (Farizi & Herwiana, 2022; Lutfi, 2023; Setyanto et al., 2026; Widodo & Belgradoputra, 2013).

2. Literature Review

Acoustic Phonetics and Mandarin Affricates

Acoustic phonetics concerns itself with the physical properties of speech sounds, particularly the temporal and spectral parameters that encode phonological contrasts (Irawan, 2017; Reetz & Jongman, 2020). For Mandarin affricates specifically, three parameters carry the primary descriptive burden in the existing literature. VOT measures the interval between stop release and the onset of voicing in the following vowel, with longer values indexing greater aspiration (Lisker & Abramson, 1964). FD measures the duration of aperiodic frication noise following the burst, serving as a secondary temporal index of aspiration energy (Forrest et al., 1988; Pollock, 2024). CoG, derived from spectral slice analysis across the fricative portion of the affricate, serves as the primary indicator of place of articulation, with lower values indexing retroflex constrictions and higher values indexing alveolar or dental contact (Gordon et al., 2002; Lee-Kim, 2014; S. Li & Gu, 2015; Narhan et al., 2023; Zhou, 2019).

Large-scale acoustic studies of native Mandarin production have established that normalized FD and amplitude are the primary perceptual and acoustic cues to the aspirated-unaspirated contrast, while spectral energy distribution is relied upon to distinguish place of articulation among the three affricates series (S. Li & Gu, 2015). Retroflex segments consistently show lower CoG than their alveolar counterparts, reflecting the longer back cavity produced by tongue retraction (Lee-Kim, 2014).

Research conducted on Taiwan and Singapore Mandarin has further demonstrated that the alveolar-retroflex distinction is subject to sociolinguistically conditioned reduction, manifesting as systematic upward shifts in CoG that bring retroflex productions closer to the alveolar range (Chang, 2012; Goh et al., 2024).

These findings establish the acoustic benchmarks against which L2 productions in the present study are evaluated. Consistent patterns of difficulty with Mandarin affricates have been documented across a range of L2 learner populations. Thai learners tend to narrow the aspirated-unaspirated contrast, producing temporal differentiation that is present but reduced to native norms (Wu & Sang, 2026). Learners with Malay and Japanese backgrounds show negative transfer patterns in which Mandarin affricates are assimilated to familiar L1 categories, observable through VOT deviations and spectral shifts in CoG (H. Li et al., 2019; Wang et al., 2013). Among Indonesian learners specifically, errors are clustered around the segments that are most typologically distant from the L1 inventory, including retroflex initials, aspirated obstruents, and the palatal affricate series, and L1 transfer combined with insufficient explicit instruction is proposed as the primary explanatory factor (D. D. Liu et al., 2021). These studies, however, were conducted in formal educational settings in which standard Pinyin was used as the orthographic medium of instruction. The role of locally adapted, non-standard romanization in vocational training contexts has not been empirically examined in any of them, and that gap is the starting point of the present investigation.

Orthographic Scaffolding, L1 Transfer, and Fossilization in Adult L2 Phonology

The relationship between orthographic input and L2 phonological representation is well established across a substantial body of acquisition research. Literate adult learners do not process spoken L2 input independently of its written form; they internalize the grapheme-phoneme mappings that their instructional materials provide and build phonological representations on the basis of those mappings (Bassetti, 2007, 2008, 2024). Experimental work on Mandarin learners demonstrates that even standard Pinyin, with its relatively transparent correspondences, induces systematic segmental deletions when its spellings underrepresent phonological segments (Bassetti, 2007). Orthographic effects of this kind have been documented across a wide range of L2 contexts, consistently showing that opaque or misleading spellings produce pronunciation errors that persist even after extended exposure to the spoken target language (Bassetti, 2006, 2008, 2024; Mok et al., 2018).

The mechanism through which orthographic input shapes phonological representation is accounted for within the theoretical framework of Flege's Speech Learning Model (SLM) (Flege, 1995; Flege & Bohn, 2021). The model proposes that L2 phonological categories are either formed as new categories or suppressed in favor of existing L1 categories, depending on the degree of perceived similarity between the L2 sound and sounds already present in the L1 inventory. When an L2 sound is perceived as sufficiently similar to an L1 sound, the learner assigns the L2 sound to the existing L1 category rather than constructing a new one, a process the model terms equivalence classification. A romanization system that represents L2 sounds through L1 graphemes does not merely reflect this perceptual tendency; it actively reinforces it by making the visual similarity between L2 and L1 sounds explicit on the page. The learner who reads "Ce" as the representation of Mandarin /zh/ is not simply making a perceptual error; she is being instructed to make it, and that instruction is repeated across every practice session in which the module is consulted.

When equivalence classification is established through repeated orthographically guided practice, the phonological pattern that results is understood within SLA theory as being susceptible to fossilization, the stabilization of non-target interlanguage forms that resist further development even under continued exposure to the target language (Z. Han, 2004; Saville-Troike & Barto, 2016). The risk of fossilization is amplified considerably when the instructional orthography is not merely underspecified, as standard Pinyin sometimes is, but is constructed from L1 graphemic material that maps systematically onto L1 phonological categories. In this situation, the learner is not working around gaps in an otherwise phonologically informed system; she is building L2 phonological representations on an L1 foundation that was laid by the curriculum itself.

Dmitrieva (2024) describes this as the mechanism through which institutional material choices determine the phonological architecture that emerges in learner production, and the concept of orthographic scaffolding from psycholinguistic research on adult L2 literacy provides a complementary way of naming it: a scaffold constructed from L1 material will produce an L2 structure that is shaped by the L1, irrespective of the learner's effort or motivation (J.-I. Han et al., 2025; Paltridge & Prior, 2024; Widad et al., 2025).

The three theoretical constructs of L1 transfer, orthographic scaffolding, and fossilization are therefore understood in this study not as independent phenomena but as components of a single causal chain. L1-based orthographic scaffolding biases the learner's phonological representations toward existing L1 categories through equivalence classification, which is the mechanism of transfer. When those representations are reinforced through repeated practice and are never challenged by corrective feedback, the transferred categories stabilize into fossilized interlanguage patterns. The present study traces this chain from its institutional origin in the training module to its acoustic endpoint in the speech of five workers in Taiwan.

Language Training Policy and the Linguistics Rights of Blue-Collar Migrants

The linguistic preparation of Indonesian migrant workers is formally regulated through the LPKS and BLK (*Balai Latihan Kerja*) system, which is overseen by Ministry of Manpower of the Republic of Indonesia (*Kementerian Ketenagakerjaan/Kemnaker*) and subject to accreditation requirements managed in part by BP2MI. In practice, however, the regulatory framework governing these institutions specifies training duration and subject coverage without attending to the phonological accuracy or pedagogical soundness of the language materials used (Allen, 2002; Purba et al., 2020). This regulatory gap creates conditions in which commercial incentives rather than linguistic criteria determine curriculum design. Training providers are placed in competition on the basis of cost and speed of certification, and the adoption of simplified romanization systems that offer the appearance of immediate readability, without regard for their phonological accuracy, is a predictable institutional response to that competitive environment (Farizi & Herwiana, 2022).

The consequences of this system fall disproportionately on workers, whose linguistic rights to accurate phonological input are not recognized as rights within the current framework (Lutfi, 2023; Setyanto et al., 2026). Research on migrant language learning more broadly has shown that blue-collar workers in destination countries operate in communicative ecologies that rarely provide the kind of corrective phonological feedback available to university learners, meaning that errors established during training tend to persist rather than self-correct through workplace exposure (Norton, 2016; Widodo & Belgradoputra, 2013). The combination of phonologically compromised pre-departure instruction and a corrective-feedback-poor workplace environment constitutes what this study treats as a structurally produced pathway to fossilization, one that originates in institutional choices rather than learner limitations.

It is within this policy landscape that the present study makes its contribution. Triangulated empirical evidence is provided that connects a specific institutional romanization practice to measurable acoustic outcomes in worker speech, thereby making traceable a causal chain that extends from a curriculum design decision in a Ponorogo training room to a spectral deviation in a Praat waveform in Taiwan. That chain, once traced, has implications for regulatory frameworks and curriculum designers that cannot reasonably be addressed by recommending that individual workers practice more carefully.

3. Method

This study was designed as an exploratory mixed-methods case study design, bounded by a single LPKS in Ponorogo, East Java, whose alumni formed a phonologically and instructionally homogeneous participant group (Creswell, 2014; Creswell & Poth, 2018). Five female PMI aged 37 to 49, all graduates of the same institution and currently employed in Taiwan, were recruited through homogeneous purposive sampling. The restriction to female participants reflects the demographic composition of the relevant LPKS cohort and reduces physiological variation in acoustic output (H. Li et al., 2019; Reranta & Huntley, 2022). Participants were recruited through snowball sampling, a non-probability technique in which initial participants refer subsequent ones through personal networks (Creswell, 2009; Miles et al., 2014). Two inclusion criteria were established.

First, participants were required to be active Indonesian migrant workers residing and employed in Taiwan at the time of data collection. Second, participants were required to have completed Mandarin language training specifically at LPKS Eka Management Cipta Sejahtera in Ponorogo, so that exposure to the training module under examination could be confirmed

for every participant. Five participants meeting both criteria were ultimately recruited, consistent with the depth requirements of in-depth case study methodology.

The restriction to female participants reflects the demographic composition of the LPKS cohort under study and the broader PMI population placed in Taiwan's domestic care and service sector (BP2MI, 2025). This restriction also reduces physiological variation in vocal tract configuration that might otherwise complicate acoustic comparisons across speakers, given that fundamental frequency and formant ranges differ systematically between male and female speakers (H. Li et al., 2019; Reranta & Huntley, 2022). Five participants are consistent with the depth requirements of in-depth case study methodology, where analytical sufficiency is determined by the richness of triangulation rather than by sample size (Creswell, 2014; Creswell & Poth, 2018).

Three data sources were collected, analyzed, and triangulated. Speech recordings were made in WAV format at sampling rate of 44,100 Hz and analyzed in Praat to extract VOT, FD, and CoG measurements for productions of /z/, /c/, /j/, /q/, /zh/, and /ch/ in citation-form target words (Boersma et al., 2024). Semi-structured interviews conducted individually with each of the five participants in Indonesian and analyzed thematically following the procedures of Braun & Clarke (2022). Interviews were designed to elicit information about participants' Mandarin learning histories, their experience with the training module, their awareness of Mandarin pronunciation as phonologically distinct from Indonesian, and the communication strategies they had developed in the Taiwanese workplace. The LPKS training module was analyzed as a pedagogical artifact, with all romanization representations of affricate initials extracted and mapped against the Indonesian grapheme-phoneme correspondences they most plausibly activate (Alwi et al., 2000; Bagaskara & Delijar, 2026; Dmitrieva, 2024).

Acoustic measurements were extracted from waveform and spectrogram displays in Praat for each target word production by each participant. VOT was measured as the duration in milliseconds between the point of stop burst release which identifiable as the moment of air pressure release on the waveform, and the onset of periodicity in the following vowel, indicated by the appearance of regular striations in the spectrogram reflecting vocal fold vibration. FD was measured as the duration of the frication segment within each affricate token. The onset of FD was marked at the emergence of high-frequency aperiodic noise following the burst, and the offset was marked at the boundary immediately preceding formant transitions into the following vowel. CoG was extracted from the mean spectrum of the fricative segment already identified for FD measurement. The Get Center of Gravity command under the Query menu in Praat was applied with a power factor (p) of 2.0, which weights the frequency distribution by the square of amplitude and yields a weighted mean frequency index reflecting the concentration of spectral energy across the fricative interval. Lower CoG values are associated with retroflex constrictions, in which tongue retraction creates a longer back cavity that concentrates energy in lower frequency bands, while higher CoG values index alveolar or dental articulations in which the shorter front cavity shifts energy concentration toward higher frequencies. Mean values and ranges per affricate category were computed and compared against native reference values from the acoustic literature (Forrest et al., 1988; Gordon et al., 2002; Lisker & Abramson, 1964; Pollock, 2024; Zhou, 2019). Given the exploratory and case-study orientation of the research, no inferential significance threshold was applied. The goal was pattern description and triangulated interpretation rather than statistical generalization, and the acoustic data are presented accordingly as evidence for phonological patterns whose origins are then traced through interview testimony and document analysis. (Lee-Kim, 2014; S. Li & Gu, 2015).

Triangulation across the three sources followed a convergent logic: acoustic patterns, interview testimony, and module analysis were examined separately and then integrated (Bagaskara, 2026). Where all three sources converged, interpretive claims were treated as well-supported; where divergence or complexity appeared, it was reported rather than resolved.

The three sources accomplish what none of them could accomplish alone, the path is traced from an orthographic decision made in a training room in Ponorogo to an acoustic measurement in a Praat spectrogram in Taiwan.

4. Results and Discussion

The findings are presented in three thematic sections, each of which brings together acoustic measurements, module evidence, and interview testimony around a distinct but interrelated dimension of the phonological problem. The first section addresses the collapse of aspiration contrasts in the alveolar and palatal series. The second addresses the systematic dentalization of retroflex affricates. The third addresses the communicative ecology in Taiwan and how it reinforces rather than corrects the interlanguage patterns established during training. This structure departs from the conventional practice of separating acoustic results from qualitative discussion, and the departure is intentional. What the data show cannot be understood apart from why it happens, and the why is only accessible when the three data sources are read against one another.

Graphemic Deception and the Collapse of Aspiration Contrasts

Before turning to individual speakers and waveform evidence, it is useful to establish the overall acoustic profile of the participant group across all six affricate categories. Table 1 summarizes the mean values and observed ranges for VOT, FD, and CoG per affricate, alongside native reference ranges drawn from the acoustic literature (Lee-Kim, 2014; S. Li & Gu, 2015). The deviations documented there are not random. They form a pattern, and that pattern has a shape that can be traced back to the orthographic decisions visible in the training module.

The acoustic profile of the five participants reveals a pattern that is neither random nor attributable to individual variation in learning capacity. What emerges is a systematic distortion of the Mandarin affricate system in a direction that is predictable from the orthographic choices embedded in the training module. Table 1 displays this pattern at the aggregate level.

Table 1. Summary of Acoustic Parameters for Mandarin Affricates.

Affricate	Type	Parameter	Native Range (literature)	PMI Range (R1–R5)	Main Deviation
/z/	alveolar, non-asp	VOT	20–40 ms	77–197 ms	Strong over-lengthening
		FD	0–20 ms	42–167 ms	Excessive frication
		CoG	4000–8000 Hz	4039–6649 Hz	Within alveolar range
/c/	alveolar, asp	VOT	80–120 ms	88–190 ms	Overlap with /z/
		FD	60–100 ms	63–155 ms	Often longer than native
		CoG	4000–8000 Hz	3961–6346 Hz	Within alveolar range
/j/	palatal, non-asp	VOT	20–40 ms	98–145 ms	Strong over-aspiration
		FD	0–20 ms	48–123 ms	Excessive frication
		CoG	3000–5000 Hz	4417–6124 Hz	Elevated; toward alveolar
/q/	palatal, asp	VOT	80–120 ms	50–185 ms	High variability
		FD	60–100 ms	20–106 ms	Sometimes shorter than /j/
		CoG	3000–5000 Hz	839–6102 Hz	Highly unstable
/zh/	retroflex, non-asp	VOT	20–40 ms	91–155 ms	Strong over-aspiration
		FD	0–20 ms	44–126 ms	Long frication
		CoG	2000–4000 Hz	4092–4591 Hz	Too high (dentalization)
/ch/	retroflex, asp	VOT	80–120 ms	87–205 ms	High variation
		FD	60–100 ms	22–157 ms	Unstable
		CoG	2000–4000 Hz	2674–6144 Hz	Extreme upper values

The most immediately striking pattern in Table 1 is that non-aspirated affricates /z/ and /j/ are being produced with VOT values that substantially exceed. That patterns not only at their native targets, but in many tokens also exceed the native ceiling for their aspirated counterpart /c/ and /q/. The aspiration contrast that is phonemically fundamental to Mandarin has effectively ceased to function in the participants' productions. When /z/ is produced with a VOT of up to 197 ms, and /c/ with a VOT that in many tokens falls within the same range, a Mandarin listener cannot use temporal cues to distinguish the two segments. The minimal pair has become an acoustic near-synonym pair.

This collapse is not, however, arbitrary. It follows directly from the romanization conventions in the LPKS module. Table 2 presents the specific grapheme mappings that are at work.

Table 2. Pinyin, LPKS Romanization, and L1 Category Mappings for Affricate Initials.

Mandarin Word	Pinyin	Target Affricate(s)	LPKS Romanization	Probable L1 Association	Projected Phonetic Effect
每次	měici	/c/ [ts ^h]	Mei Jhe	“jh” ≈ /dʒ/	Aspiration replaced by voiced fricative quality
自己	zìjǐ	/z/ [ts]	Ce Ci	“c” ≈ /tʃ/	/z/ assimilated to Indonesian /tʃ/
汽车	qìchē	/q/ [tɕ ^h] /ch/ [tɕ ^h]	Jhi Jhe	“jhe” ≈ /dʒ/	Palatal and retroflex merged toward /dʒ/
知道	zhīdào	/zh/ [tʂ]	Ce Tao	“ce” ≈ /tʃ/	Retroflex replaced by alveolar /tʃ/ quality

Source: LPKS (n.d.); Alwi et al. (2000); H. Li et al. (2019).

The logic of the module’s orthographic system becomes clear in Table 2. The letter “c” in Indonesian orthography represents the phoneme /tʃ/, a voiced palato-alveolar affricate that carries no aspiration distinction. When *zìjǐ* is printed as “Ce Ci,” a learner trained in Indonesian literacy will interpret both syllables through the lens of /tʃ/, producing a segment that in its temporal structure resembles neither the unaspirated /z/ nor its aspirated counterpart /c/ but instead a heavily fricated approximation of Indonesian /tʃ/ with extended VOT and FD. The acoustic measurements in Table 1 are exactly what one would predict from this orthographic logic.

The same logic extends to the palatal series. The grapheme cluster “jh” does not exist in standard Indonesian orthography, but its closest associative neighbor is /dʒ/, the voiced palato-alveolar affricate represented by the letter “j” in Indonesian. When *měici* is printed as “Mei Jhe,” the learner is invited to treat /c/ as though it were a voiced, unaspirated affricate, which produces exactly the kind of temporal instability visible in the /q/ column of Table 1, where VOT ranges from 50 to 185 ms with no consistent target. The speaker is not converging on a category; she is searching for one that the module has not provided. What Figure 1 follows is a Praat-based illustration of this collapse at the individual speaker level.

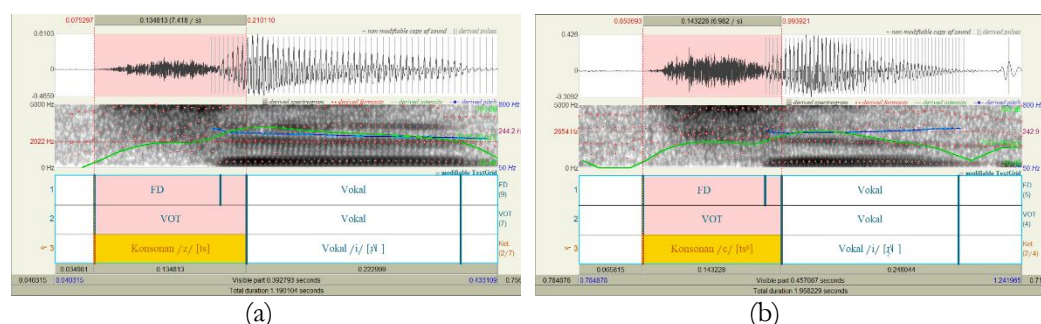


Figure 1. Praat waveform, spectrogram, and spectral slice of /z/ (a) and /c/ (b) produced by participant R2.

Figure 1 makes visible what the numbers in Table 1 described. The two productions are almost indistinguishable in their temporal envelopes. Both show a clearly defined stop closure, a burst, and then an extended interval of aperiodic frication noise that persists well beyond any native target for either aspirated or unaspirated affricates. The VOT values of 134 ms and 143 ms differ by only 9 ms, which is well within the range of normal phonetic variation and cannot reasonably be expected to function as a phonemic cue for a Mandarin listener. In practical terms, R2 is producing two tokens of the same sound where Mandarin requires two phonemically distinct sounds.

The interview data from R2 makes the pathway from orthography to production explicit. When asked how she learned to pronounce *zìjǐ*, she stated that she read it the way it was written in the book, as “Ce Ci,” and that her tongue moved automatically as though she were speaking Indonesian. This is not a description of confusion or uncertainty; it is a description of successful retrieval from an internalized phonological representation built on L1 categories.

The module taught her to read “Ce” as /tʃ/, she learned that lesson well, and her VOT values reflect that learning with precision. The error is not a performance failure; it is a competence failure that was produced by instruction.

The palatal series shows a related but distinct pattern of disruption. Where the alveolar series collapses into mutual overlap, the palatal series collapses into mutual instability, see Figure 2.

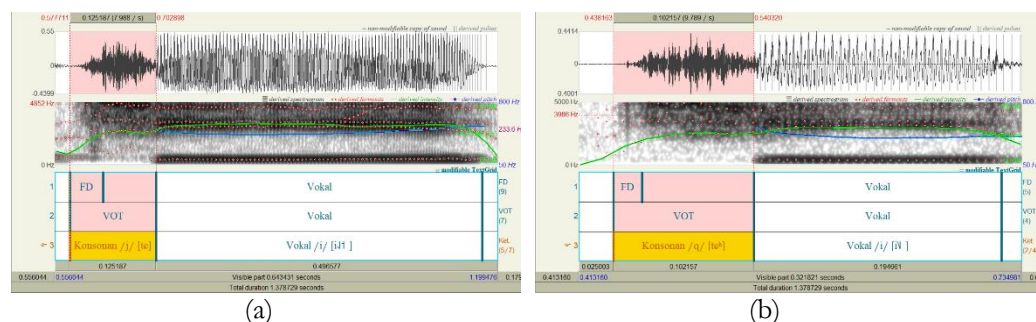


Figure 2. Praat waveform, spectrogram, and spectral slice of /j/ (a) and /q/ (b) produced by participant R4.

Figure 2 illustrates what might be called an aspiration inversion. The non-aspirated segment /j/ is produced with greater aspiration than the aspirated segment /q/. This is not a pattern that can be explained by L1 transfer alone, since Indonesian lacks both segments and therefore provides no direct phonological template for either. What the module provides instead is a unified template: both /j/ and /q/ are represented through "jh"-type spellings that activate the Indonesian /dʒ/ category. Since /dʒ/ carries no aspiration distinction in Indonesian, the learner has no orthographic basis for differentiating the two Mandarin segments, and the result is the high variability visible in the /q/ column of Table 1. The VOT for /q/ ranges from 50 to 185 ms across the five participants, a 135 ms spread that is not random scatter but reflects the absence of a stable phonological target. R4's interview comment captures this precisely: she was taught the "jh" spelling for both sounds, she never received instruction on how to differentiate them, and in the workplace she simply tries to make the sound "stronger" when she is not understood. Making a sound stronger, without a clear articulatory target, is not a path toward the native aspiration contrast; it is a path toward hyperarticulation along an L1 trajectory.

Institutionalized Dentalization and the Cost of Immediate Readability

The retroflex series presents a different kind of phonological damage, one that is visible not in the temporal domain but in the spectral domain. CoG values for /zh/ and /ch/ are the critical evidence here, and they tell a consistent story: the participants are not producing retroflex affricates. They are producing alveolar or dental affricates and labeling them with Mandarin-sounding words.

Native Mandarin retroflex affricates are characterized by low CoG values, typically in the range of 2000 to 4000 Hz, which reflects the long back cavity created by tongue retraction during retroflex articulation (Lee-Kim, 2014). The participants in this study show CoG values for /zh/ that range from 4092 to 4591 Hz, and for /ch/ that extend in some tokens up to 6144 Hz. These values are not marginally elevated; they are categorically in the wrong spectral region. A CoG of 6144 Hz belongs to the acoustic territory of dental or alveolar sibilants, not retroflexes.

The module provides a direct explanation. As shown in Table 2, /zh/ is represented in the LPKS materials as "Ce" a spelling that activates Indonesian /tʃ/, which is an alveolar affricate produced with the tongue blade at or near the alveolar ridge, not curled back toward the hard palate. The learner who internalizes "Ce" as the guide for /zh/ will produce the tongue position associated with Indonesian /tʃ/, and that tongue position yields the high CoG values observed in the data. The dentalization is not phonetic drift; it is phonetic substitution, and it was instructed. The data are consistent with an interpretation in which dentalization represents phonetic substitution that was systematically instructed through the module's romanization choices.

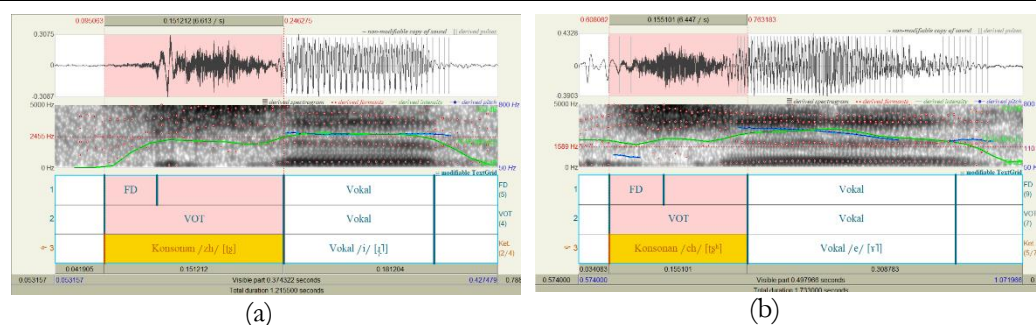


Figure 3. Praat waveform, spectrogram, and spectral slice of /zh/ (a) and /ch/ (b) produced by participant R3.

Figure 3 makes the spectral shift unmistakable. Where a native retroflex affricate would show its energy concentrated in the lower frequency bands of the spectrogram, with a spectral slice peak well below 4000 Hz, R3's productions show energy concentrated in the upper frequency bands, with the spectral slice revealing a CoG that in the /ch/ token reaches 6144 Hz. The /ch/ CoG values represent a systematic shift from retroflex to alveolar/dental articulation, a shift that is traceable to the module's representation of /zh/ as 'Ce'. This is the acoustic signature of a tongue that is positioned at the alveolar ridge, not curled back. The long back cavity that gives retroflex sounds their characteristically low spectral gravity is simply not being created.

The interview data from R3 and other participants on this point is remarkably consistent. All five participants reported that their LPKS instructors taught them to read the romanized spellings and to practice saying them quickly, but none reported receiving any instruction on tongue position, on what retroflexion means, or on how /zh/ and /ch/ differ from /z/ and /c/ at the articulatory level. One participant described the instruction in terms that might translate roughly as: "We were just told to memorize the sound next to the letter. Nobody ever explained where the tongue should go." Another participant noted that she had been using "Jhe" as her guide for /zh/ for over three years and had never been told by an employer or a native speaker that anything was wrong with her pronunciation. The communicative success she experienced was real, which means she could make herself understood in context, but it was achieved at the cost of phonemic accuracy that is in principle required for intelligibility in unfamiliar contexts or with less accommodating interlocutors (Goh et al., 2024; Widad et al., 2025).

The broader implication of the dentalization pattern points directly to the curriculum design decision that produced it. Retroflex consonants require articulatory knowledge that cannot be conveyed by a romanized spelling alone, and certainly not by a romanized spelling drawn from an orthographic system that contains no retroflex sounds. The choice to represent /zh/ as "Ce" and /ch/ as "Jhe" was almost certainly made in the interest of immediate readability — a learner who knows Indonesian orthography can immediately pronounce something when she sees "Ce," whereas a learner confronted with *zh* has no immediate decoding strategy without explicit instruction. The LPKS design prioritizes the experience of reading fluency in the short term at the expense of phonological accuracy in the long term. This is the cost of what might be called the pedagogy of instant legibility, and it is a cost borne entirely by the learner (Allen, 2002; Farizi & Herwiana, 2022; Lutfi, 2023).

Communication Ecology, Hyperarticulation Traps, and the Fossilization of Error

The phonological patterns would not necessarily persist if the participants' communicative environment in Taiwan provided corrective pressure sufficient to challenge them. That environment does not. The interview data across all five participants converges on a description of a workplace in which Mandarin pronunciation is evaluated primarily for functional sufficiency rather than phonemic precision. Employers and household members accommodate the workers' speech, rely on contextual cues to resolve ambiguity, and rarely identify specific pronunciation features as problematic. One participant described her experience of three years of daily Mandarin use in Taiwan without a single instance of an employer correcting her pronunciation, even when she knew from the listener's expression that something had not been understood immediately. The correction, when it came at all, was at the level of meaning, not sound.

Before examining how this workplace ecology reinforces phonological patterns, it is necessary to distinguish clearly between two mechanisms operating at different points in time. The phonological patterns are products of the training institution, where L1-based orthography shapes phonological category formation through equivalence classification during the training phase. The persistence of these patterns, however, is a product of the communicative ecology in Taiwan, where functional sufficiency is achieved through accommodation and contextual inference rather than through corrective feedback on segmental accuracy. Disentangling these influences is essential for understanding why fossilization occurs. The institutional phase installs phonologically incorrect categories, and the workplace phase prevents those categories from being challenged or dismantled. Neither phase alone would produce fossilization; it is their sequential combination that makes fossilization structurally inevitable.

This communicative ecology has two consequences for phonological development, both of which are visible in the data. The first is that the interlanguage patterns established during training are never challenged by external feedback, and they therefore calcify into what the second language acquisition literature describes as fossilization (Z. Han, 2004; Saviile-Troike & Barto, 2016). Fossilization in this study is not merely theoretical; it is evidenced by the fact that participants who had been living and working in Taiwan for between two and four years still showed the same systematic VOT over-extension and CoG elevation that would be expected from learners in the very early stages of acquisition. Time and exposure had not moved them toward native norms, a finding consistent with the hypothesis that corrective feedback, rather than mere exposure to input, drives phonological development in adult L2 learners.

The second consequence is that the participants had independently developed compensatory strategies that felt to them like efforts at self-correction but were acoustically counterproductive. All five participants described a practice of exaggerating pronunciation when they felt misunderstood, making sounds louder, longer, or clearer. This behavior is consistent with what the phonetics literature describes as hyperarticulation, the enhancement of acoustic cues in response to perceived communicative difficulty (Chang, 2012; Reranta & Huntley, 2022). Under normal circumstances, hyperarticulation can be a functional strategy that enhances the salience of contrastive acoustic features. When a native speaker hyperarticulates, she typically amplifies the very features that are phonemically relevant. When a learner with a fossilized interlanguage hyperarticulates, she amplifies the features of her interlanguage categories, which may bear little relationship to the phonemically relevant features of the target language. In the present data, hyperarticulation appears to be driving VOT and FD values upward without moving CoG values in the direction of retroflex production. A participant who makes /zh/ louder and longer is extending the duration of the frication interval of her alveolar /tʃ/ approximation, not curling her tongue back toward retroflex position. The acoustic output has more energy, but it has more energy in the wrong spectral location. This is the hyperarticulation trap, a strategy that feels like improvement but is quantifiably a magnification of error (Galante & Piccardo, 2022; J. Liu & Jongman, 2013; Norton, 2016).

The implications of this trap extend beyond individual participants. If a training system installs phonologically incorrect categories, and if the communicative environment provides no corrective feedback, and if the learner's only available self-correction strategy is to amplify what she already knows how to do, then the fossilization of error is not merely possible; it is structurally guaranteed. The workers in this study did not arrive at their interlanguage affricate system through carelessness or insufficient effort. They arrived there by following the instructions they were given in a training module that was constructed under conditions of commercial pressure, in an environment that subsequently gave them no reason to doubt those instructions (Flege & Bohn, 2021; J.-I. Han et al., 2025).

This finding carries a direct implication for how the phenomenon should be framed at the policy and practitioner level. The phonological failures documented in this study are not individual learning failures. They are the predictable output of a curriculum that was designed under economic and time pressure to produce the appearance of communicative readiness. Recommending to individual workers that they practice more carefully or listen more attentively is not a meaningful response to a systemic institutional failure.

The appropriate response is institutional, and it operates at multiple levels by regulatory oversight of LPKS materials, pedagogical transformation within training institutions, and structured early intervention in the workplace before fossilization is complete (Allen, 2002; Farizi & Herwiana, 2022; Setyanto et al., 2026).

5. Conclusion

This study has traced a causal chain from a specific set of orthographic choices in an LPKS training module to measurable phonological deviations in the speech of five Indonesian migrant workers in Taiwan. Non-aspirated Mandarin affricates were produced with VOT values far exceeding native ceilings, collapsing the aspiration contrast that is phonemically fundamental to the language. Retroflex affricates were replaced by alveolar or dental productions, as evidenced by CoG values that in some tokens reached 6144 Hz. These values placed the productions squarely in the spectral territory of alveolar sibilants rather than retroflexes. These patterns were traceable, through triangulation of acoustic data, interview testimony, and module analysis, to the systematic L1-based grapheme-phoneme mappings encoded in the training materials. While the case study design with $n=5$ does not permit generalization beyond these five participants, the consistency of the pattern across all participants and across multiple acoustic parameters suggests a systematic institutional effect rather than idiosyncratic variation. They were sustained by a Taiwanese workplace ecology that provided functional communicative sufficiency without phonemic corrective pressure, and they were reinforced by hyperarticulation strategies that amplified interlanguage categories rather than target-like ones. The central argument of this study is that these outcomes are institutional rather than individual. The phonological failures documented here were not produced by insufficient learner effort; they were produced by a curriculum designed to prioritize immediate readability over phonological accuracy, in a regulatory environment that imposes no standards on the phonological content of LPKS teaching materials.

Three policy recommendations follow from this argument, directed at Kemnaker and BP2MI as the regulatory authorities with jurisdiction over LPKS accreditation. First, a national phonological standard for foreign language teaching materials in LPKS and BLK settings should be established as a condition of institutional accreditation, requiring that romanization systems demonstrably avoid mapping Mandarin phonological categories onto L1 graphemes that encode categorically different sounds. Second, explicit articulatory instruction including tongue position description, native speaker audio modeling, and where resources permit Praat-based visual feedback, should be mandated as a core component of Mandarin pronunciation training rather than an optional supplement. Third, structured pronunciation feedback mechanisms should be integrated into both the training phase and the early workplace placement phase, before fossilization of non-target patterns is complete.

Beyond policy-level intervention, the findings carry implications for practitioners who work directly with migrant learners. For LPKS instructors, the evidence suggests that reliance on L1-based orthography as a primary teaching tool is pedagogically counterproductive, and that allocating class time to articulatory description and visual feedback through Praat is not a luxury but a necessity for accurate phonological category formation. For curriculum designers, the choice between Hanyu Pinyin and locally adapted romanization is not merely a matter of convenience; it determines the shape of the phonological system that emerges in learner production. For placement officers in Taiwan who manage early interaction between newly arrived PMI and their employers, structured feedback on pronunciation in the first weeks of employment, before interlanguage patterns fossilize, offers the most cost-effective intervention point. These practitioners need not wait for national regulation to begin implementing changes informed by the evidence in this study.

Future research should examine whether the patterns documented here replicate across other LPKS institutions using locally adapted romanization, and whether controlled pedagogical interventions replacing non-standard romanization with Pinyin produce measurable acoustic improvement in pre-departure learner cohorts. What this study has established is that the path from an orthographic decision in a Ponorogo classroom to an acoustic deviation in a Taiwanese workplace is empirically traceable, and that the institutions responsible for that decision bear a form of pedagogical accountability that the current

regulatory framework does not yet require them to honor. Correcting this is not only a matter of educational quality. It is a matter of linguistic justice.

Data Availability Statement: The data supporting the findings of this study, including raw acoustic measurement files, Praat output, and interview transcripts, are not publicly archived due to the privacy and confidentiality considerations applicable to the participant group. Researchers who wish to access the dataset for verification or scholarly purposes are invited to contact the corresponding author directly at the email address provided on the title page. Requests will be considered on a case-by-case basis in accordance with applicable ethical guidelines for research involving human participants.

Acknowledgments: The authors wish to express sincere gratitude to the five participants who generously shared their time, their voices, and their experiences as the foundation of this study. Their willingness to speak openly about the difficulties of learning and working in a language they were given insufficient tools to master made this research possible in the most fundamental sense. Appreciation is extended to LPKS Eka Management Cipta Sejahtera in Ponorogo, East Java, for permitting access to the training module examined in this study. The institution's cooperation was essential to the document analysis component of the triangulation design. The authors are grateful to the Faculty of Cultural Studies, Universitas Brawijaya, for institutional support during the preparation of this manuscript. No external funding was received for this research. The authors declare no conflict of interest.

References

- Afrina, U., & Cleveresty, T. B. (2021). Analisis Penggunaan Hanyu Pinyin dalam Pelafalan Bahasa Mandarin Berdasarkan Unsur Pembentuknya. *Metahumaniora*, 11(2), 135–144. <https://doi.org/10.24198/metahumaniora.v11i2.32776>
- Allen, L. Q. (2002). Teachers' Pedagogical Beliefs and the Standards for Foreign Language Learning. *Foreign Language Annals*, 35(5), 518–529. <https://doi.org/10.1111/j.1944-9720.2002.tb02720.x>
- Alwi, H., Moeliono, A. M., Dardjowidjojo, S., & Lapoliwa, H. (2000). *Tata bahasa baku bahasa Indonesia* (3rd ed.). Balai Pustaka.
- Bagaskara, A. S. D. (2026). *Produksi fonetik konsonan afrikat bahasa Mandarin pekerja migran Indonesia di Taiwan: Studi kasus LPKS Eka Management Cipta Sejahtera* [Unpublished thesis]. Universitas Brawijaya.
- Bagaskara, A. S. D., & Delijar, R. M. (2026). The Romanization-Induced Phonetic Deviation in Mandarin Affricates of Indonesian Migrant Workers' Interlanguage in Taiwan [Submission manuscript]. *Pioneer: Journal of Language and Literature*.
- Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language & Literacy*, 9(1), 95–114. <https://doi.org/10.1075/wll.9.1.07bas>
- Bassetti, B. (2007). Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language. In A. Guder, X. Jiang, & Wan. Y (Eds.), *The Cognition, Learning and Teaching of Chinese Characters*. Beijing Language and Culture University Press. <https://eprints.bbk.ac.uk/id/eprint/528>
- Bassetti, B. (2008). Orthographic Input and Second Language Phonology. In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 191–206). Multilingual Matters. <https://doi.org/10.21832/9781847691118-013>
- Bassetti, B. (2024). Orthographic Effects in the Phonetics and Phonology of Second Language Learners and Users. In *The Cambridge Handbook of Bilingual Phonetics and Phonology* (pp. 699–720). Cambridge University Press.
- Boersma, P., Weenink, D., & Shchupak, A. (2024). *Praat: Doing Phonetics by Computer: Version 6.4.67* (Version 6.4.67). <https://praat.org>
- BP2MI. (2025). *Laporan Tahunan Publikasi Data Pekerja Migran Indonesia Tahun 2024*. https://kp2mi.go.id/uploads/statistik/images/data_06-02-2025_Laporan_Tahunan__Publikasi_Data_PMI_2024.pdf
- Braun, Virginia., & Clarke, Victoria. (2022). *Thematic Analysis: a Practical Guide*. SAGE.

- Chang, Y.-H. (2012). *Variability in cross-dialectal production and perception of contrasting phonemes: The case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin* [Dissertation, University of Illinois]. <https://www.ideals.illinois.edu/items/42188>
- Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). SAGE.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design* (4th ed.). SAGE.
- Dmitrieva, O. (2024). The Phonetics and Phonology of Adult L2 Learners in the Classroom. In M. Amengual (Ed.), *The Cambridge Handbook of Bilingual Phonetics and Phonology* (pp. 521–541). Cambridge University Press. <https://doi.org/10.1017/9781009105767.024>
- Farizi, F., & Herwiana, S. (2022). A Study of Teaching Speaking Techniques Used at Peace English Course Pare, Kediri. *Pioneer: Journal of Language and Literature*, 14(1), 43. <https://doi.org/10.36841/pioneer.v14i1.1382>
- Fauzi, I. (2021). The Variability in Phonology of Indonesian Learner's Interlanguage: A case study on English marked-fricatives. *International Journal of Language Education*, 5(4), 285–295. <https://doi.org/10.26858/ijole.v5i4.19468>
- Fitriyah, Z., & Subandi, S. (2020). Analisis Kesalahan Pelafalan Huruf Konsonan (声母 shēngmǔ) Oleh Para Tokoh dalam Drama Thailand yang Berjudul Love Beyond Frontier. *Jurnal Bahasa Mandarin UNESA*, 4(2). <https://ejournal.unesa.ac.id/index.php/manadarin/article/view/45181>
- Flege, J. E. (1995). Second Language Speech Learning Theory, Findings, and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233–277). York Press.
- Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901>
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America*, 84(1), 115–123. <https://doi.org/10.1121/1.396977>
- Galante, A., & Piccardo, E. (2022). Teaching pronunciation: toward intelligibility and comprehensibility. *ELT Journal*, 76(3), 375–386. <https://doi.org/10.1093/elt/ccab060>
- Goh, H. L., Woon, F. T., Moisik, S. R., & Styles, S. J. (2024). Contrastive Alveolar/Retroflex Phonemes in Singapore Mandarin Bilinguals: Comprehension Rates for Articulations in Different Accents, and Acoustic Analysis of Productions. *Language and Speech*, 67(4), 924–944. <https://doi.org/10.1177/00238309231205012>
- Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32(2), 141–174. <https://doi.org/10.1017/S0025100302001020>
- Han, J.-I., Kim, S. Y., & Kim, J.-Y. (2025). Effects of Orthographic Input and Inhibitory Control on Second-Language Speech Production. *Language and Speech*, 68(2), 391–412. <https://doi.org/10.1177/00238309241270737>
- Han, Z. (2004). *Fossilization in adult second language acquisition* (D. Singleton, Ed.; 1st ed.). Multilingual Matters.
- Irawan, Y. (2017). *Fonetik Akustik: Sebuah Pengantar Telaah Wujud Akustik Bahasa* (Vol. 1). Angkasa.
- Lee-Kim, S.-I. (2014). Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association*, 44(3), 261–282. <https://doi.org/10.1017/S0025100314000267>
- Li, H., Baryadi, I. P., & Wijana, I. D. P. (2019). Sound Pattern of Indonesian Plosives. *Linguistik Indonesia*, 37(1), 1–12. <https://doi.org/10.26499/li.v37i1.84>

- Li, J., Xu, R. B., & Mok, P. P. K. (2014). Effect of Orthography on L2 Production of Mandarin Tones. *Proc. 4th International Symposium on Tonal Aspects of Languages (TAL 2014)*, 147–151.
- Li, S., & Gu, W. (2015). Acoustic Analysis of Mandarin Affricates. *INTERSPEECH*, 1680–1684. <https://doi.org/10.21437/Interspeech.2015>
- Li, X.-H., Yang, D.-S., Jing, J., Zheng, J., Luo, D., Wang, X., Li, Y.-Y., Huang, M.-Y., & Zhang, X.-L. (2011). Mandarin phonological process in mandarin-spoken university students and cantonese spoken university students. *Chinese Mental Health Journal*, 25(7), 528–532. https://caod.oriprobe.com/articles/27543868/Mandarin_phonological_process_in_mandarin_spoken_u.htm
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *WORD*, 20(3), 384–422. <https://doi.org/10.1080/00437956.1964.11659830>
- Liu, D. D., Artawa, K., Budiarsa, M., Suastra, I. M., & Septevany, E. (2021). Pronunciation Errors of Indonesian Mandarin Learners. *TIJOLAC: The International Journal of Language and Cultural*, 3(1), 66–82.
- Liu, D., & Reed, M. (2021). Exploring the Complexity of the L2 Intonation System: An Acoustic and Eye-Tracking Study. *Frontiers in Communication*, 6. <https://doi.org/10.3389/fcomm.2021.627316>
- Liu, J., & Jongman, A. (2013). American Chinese learners' acquisition of L2 Chinese affricates /ts/ and /tsh/. *Proceedings of Meetings on Acoustics*, 060005. <https://doi.org/10.1121/1.4798223>
- LPKS. (n.d.). *Modul Bahasa Mandarin*. LPKS Eka Management Cipta Sejahtera.
- Lutfi, R. S. (2023). The Impact of English Language on The Existence of Indonesian Language in The Age of Globalization. *LITERA: Jurnal Bahasa Dan Sastra*, 9(2), 9–14. <https://doi.org/10.36002/litera.v9i2.3087>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: a methods sourcebook* (3rd ed.). SAGE Publications, Inc.
- Mok, P. P. K., Lee, A., Li, J. J., & Xu, R. B. (2018). Orthographic effects on the perception and production of L2 mandarin tones. *Speech Communication*, 101, 1–10. <https://doi.org/10.1016/j.specom.2018.05.002>
- Narhan, R., Sholihatun, P., & Syarfina, T. (2023). Analisis Frekuensi, Intensitas, dan Durasi pada Bahasa Turki oleh Native Speaker dan Non-Native Speaker menggunakan Praat. *LINGUA: Jurnal Bahasa, Sastra, Dan Pengajarannya*, 20(2), 351–372. <https://doi.org/10.30957/lingua.v20i2.840>
- Norton, B. (2016). Identity and Language Learning: Back to the Future. *TESOL Quarterly*, 50(2), 475–479. <https://doi.org/10.1002/tesq.293>
- Paltridge, B., & Prior, M. T. (2024). *The Routledge Handbook of Second Language Acquisition and Discourse*. Routledge. <https://doi.org/10.4324/9781003177579>
- Pollock, M. (2024). Buenas no[tʃ]es y mu[ts]isimas gracias: A sociophonetic study of the alveolar affricate in peninsular Spanish political speech. *Languages*, 9(6), 218. <https://doi.org/10.3390/languages9060218>
- Purba, Y. A., Aini, Y. N., Asiati, D., & Ngadi, N. (2020). Kesenjangan Kepesertaan Jaminan Sosial Tenaga Kerja di Indonesia: Analisis Data Sakernas 2018. *Jurnal Ketenagakerjaan*, 15(2). <https://doi.org/10.47198/naker.v15i2.72>
- Reetz, Henning., & Jongman, Allard. (2020). *Phonetics: transcription, production, acoustics, and perception* (2nd ed.). John Wiley & Sons, Inc.
- Reranta, R. C., & Huntley, M. L. (2022). Acoustic Parameters Giving the Angry Impression in Lampungnese Neutral Speech. *Linguistik Indonesia*, 40(1), 49–61. <https://doi.org/10.26499/li.v40i1.272>
- Saville-Troike, Muriel., & Barto, Karen. (2016). *Introducing second language acquisition* (3rd ed.). Cambridge University Press.

- Setyanto, A., Nishimura, A., & Fernando, H. (2026). *Japanese as a Second Language for Indonesian Migrant Worker Candidates: Linguistic and Intercultural Challenges in Pre-departure Training*. <https://doi.org/10.21203/rs.3.rs-8857635/v1>
- Thamrin, L., Veronica, T., Lusi, L., & Suhardi, S. (2022). Problem Identification of Learning Chinese Proverbs at Faculty of Teacher Training and Education Tanjungpura University. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, 8(3), 716. <https://doi.org/10.33394/jk.v8i3.5671>
- Wang, M., Lin, C. Y., & Yang, C. (2013). Contributions of Phonology, Orthography, and Morphology in Chinese-English Bilingual Acquisition: A One-Year Longitudinal Study. In X. Chen, Q. Wang, & Y. C. Luo (Eds.), *Reading Development and Difficulties in Monolingual and Bilingual Chinese Children* (pp. 191–211). Springer. https://doi.org/10.1007/978-94-007-7380-6_10
- Widad, S., Maulindah, D., Ramdhani, A. N. F., & M. Yunus Abu Bakar. (2025). Dinamika Pemerolehan dan Pembelajaran Bahasa: Analisis Prinsip, Perbedaan Usia, dan Faktor Pendukung. *Bhinneka: Jurnal Bintang Pendidikan Dan Bahasa*, 3(3), 165–186. <https://doi.org/10.59024/bhinneka.v3i3.1481>
- Widodo, H., & Belgradoputra, R. J. (2013). Perlindungan Pekerja Migran Indonesia. *Binamulia Hukum*, 8(1).
- Wu, S., & Sang, Y. (2026). Production of Mandarin affricates by Thai learners: An acoustic study. *Acoustical Science and Technology*, 47(3), e25.72. <https://doi.org/10.1250/ast.e25.72>
- Zhou, Z. (2019). *Phonetic Evidence for a Feed-forward Model: Rounding and Center of Gravity of English [f]* [Master thesis, UCLA]. <https://escholarship.org/uc/item/11x2j8k8>